Report Documentation Page					Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.						
1. REPORT DATE **2007**		2. REPORT TYPE			3. DATES COVERED **00-00-2007 to 00-00-2007**	
4. TITLE AND SUBTITLE **Subjective Probability Distribution Elicitation in Cost Risk Analysis. A Review**					5a. CONTRACT NUMBER	
					5b. GRANT NUMBER	
					5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)					5d. PROJECT NUMBER	
					5e. TASK NUMBER	
					5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Rand Corporation,1776 Main Street,PO Box 2138,Santa Monica,CA,90407-2138**					8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)					10. SPONSOR/MONITOR'S ACRONYM(S)	
					11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT						
15. SUBJECT TERMS						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**37**		

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

TECHNICAL
REPORT

Subjective Probability Distribution Elicitation in Cost Risk Analysis

A Review

Lionel A. Galway

Prepared for the United States Air Force

Approved for public release; distribution unlimited

RAND PROJECT AIR FORCE

The research described in this report was sponsored by the United States Air Force under Contract FA7014-06-C-0001. Further information may be obtained from the Strategic Planning Division, Directorate of Plans, Hq USAF.

Library of Congress Cataloging-in-Publication Data

Galway, Lionel A., 1950-
Subjective probability distribution elicitation in cost risk analysis : a review / Lionel A. Galway.
p. cm.
Includes bibliographical references.
ISBN 978-0-8330-4011-4 (pbk. : alk. paper)
1. United States. Air Force—Appropriations and expenditures. 2. United States. Air Force—Costs. 3. United States. Air Force—Cost control. I. Title.

UG633.2.G55 2007
358.4'1622973—dc22

2007014086

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

Published 2007 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Preface

A cost estimate for a project, such as the development of a new aircraft or satellite system, carries with it an inherent probability that the actual cost will exceed the estimate—that changes in requirements, technology, the economic environment, and a multitude of other factors over the life of an acquisition project will change the final cost. One major approach to cost risk analysis—evaluating and quantifying the uncertainty of a cost estimate—has been probabilistic: expressing the uncertainty in a cost estimate as a probability distribution over a range of potential costs. Cost analysts in industry and government and researchers in statistics and management have often proposed that, to get probability distributions for platforms using new and untried technologies, expert judgment should be tapped and subjective probability distributions elicited from the experts to represent cost uncertainty. Procedures for eliciting subjective probability distributions in cost risk analysis are reviewed in this technical report.

This review of elicitation procedures is the product of two projects. The first was a RAND Corporation Internal Research and Development (IR&D) project titled "Risk Management and Risk Analysis for Complex Projects: Developing a Research Agenda," conducted in 2001–2002. Support for RAND's continuing program of self-sponsored independent research is provided, in part, by donors and by the independent research and development provisions of RAND's contracts for the operation of its U.S. Department of Defense federally funded research and development centers. The principal investigator for this research was the present author. Related research is contained in the following document:

- *Quantitative Risk Analysis for Project Management: A Critical Review*, Lionel A. Galway (WR-112-RC). This working paper reviews the literature on applying quantitative risk analytic methods to key parameters of project management—primarily, cost estimation and scheduling.

The second project for which a review of material specific to cost risk analysis was done is RAND Project AIR FORCE's "The Cost of Future Military Aircraft: Historical Cost Estimating Relationships and Cost Reduction Initiatives." The purpose of the project is to improve the tools used to estimate the costs of future weapon systems. It focuses on how recent technical, management, and government policy changes affect cost. A related document is

- *Impossible Certainty: Cost Risk Analysis for Air Force Systems*, Mark V. Arena, Obaid Younossi, Lionel A. Galway, Bernard Fox, John C. Graser, Jerry M. Sollinger, Felicia Wu,

and Carolyn Wong (MG-415-AF). This report sets out guidelines for the Air Force to consistently apply cost risk analysis to cost estimates made for new Air Force systems.

This report examines cost risk analysis methods and recommends practice and policy changes to utilize subjective opinion for improving the quantification and use of uncertainty in cost estimation. It should be of interest to cost analysis professionals who wish to quantify uncertainty when using expert opinions in cost risk analysis.

This project was conducted within the RAND Project AIR FORCE Resource Management Program and was sponsored by the Principal Deputy, Office of the Assistant Secretary of the Air Force (Acquisition), Lt Gen John D. W. Corley at the beginning of the project, and now Lt Gen Donald J. Hoffman. The project technical monitor was Jay Jordan, Technical Director of the Air Force Cost Analysis Agency. The RAND project leaders were Obaid Younossi and Mark V. Arena.

This material was included in a different form as Chapter Eleven in the *Space Systems Cost Analysis Group [SSCAG] Space Systems Cost Risk Handbook*, published in 2005 by the SSCAG and edited by Timothy P. Anderson and Raymond P. Covert. Support for writing that chapter was provided by RAND Statistics Group Methodology funding.

RAND Project AIR FORCE

RAND Project AIR FORCE (PAF), a division of the RAND Corporation, is the U.S. Air Force's federally funded research and development center for studies and analyses. PAF provides the Air Force with independent analyses of policy alternatives affecting the development, employment, combat readiness, and support of current and future aerospace forces. Research is conducted in four programs: Aerospace Force Development; Manpower, Personnel, and Training; Resource Management; and Strategy and Doctrine.

Additional information about PAF is available on our Web site at:
http://www.rand.org/paf/

Contents

Figure

Summary

It has become a truism of cost estimation that there is no "right" cost estimate; for any particular estimate, there is some inherent probability that the actual cost will exceed the estimate. A cost estimate is actually a forecast in which inherent uncertainties arise from changes in requirements, technology, the economic environment, and a multitude of other factors. One major approach to *cost risk analysis*—evaluating and quantifying the uncertainty in a cost estimate—has been probabilistic: to express the uncertainty in a cost estimate as a probability distribution over a range of possible costs. To get probability distributions for the characteristics and costs of new and untried technologies for which little data are available, cost analysts have often proposed tapping the resources of expert judgment and eliciting subjective probability distributions to quantify cost uncertainty.

Elicitation of subjective probability distributions as an area of research in its own right arose out of several developments in statistics and psychology and in demands from the field of general risk analysis. Experiments in elicitation indicated that human beings were subject to a number of serious biases, such as overconfidence in their ability to quantify uncertainty, which distorted their judgment about uncertainty.

How elicitation for cost risk purposes is actually practiced is difficult to determine. There is little information in the professional cost risk literature, and elicitation practices are very diverse—even in such disciplines as space systems acquisition, which involves a limited number of organizations and so is an area in which one might expect a high degree of methodological standardization. This, despite tutorial materials routinely recommending elicitation of expert judgment when data are scarce or when historical data might be irrelevant.

What practical advice can be given? A start can be made with the following procedure, which is synthesized from a number of sources:

- Use multiple independent experts.
- Ask an expert to provide, at a minimum, upper, lower, and most-likely values for cost elements under consideration.
- Fit a triangle distribution[1] to these three numbers, but using the upper and lower values to bound 90 percent of the probability (where reasonable, to counteract known biases in elicitation).

1 For the triangle distribution, the probability is set to zero outside the endpoints, while between the endpoints the density rises linearly from the lower value to the most-likely value and then decreases linearly from the most-likely value to the

- In addition to the upper, lower, and most-likely values, elicit at least two more percentiles (perhaps the 25th and 75th, as recommended by several current authors).
- Provide feedback to the expert about the results of the elicitation, including the final range of nonzero probabilities, the median estimated cost, the probability that the final cost will exceed the most-likely cost, and so on.
- Carefully document the process and the results and archive the data obtained, for future retrospective studies.

The cost-estimation community is in general agreement that probabilistic methods of quantifying and reasoning with uncertainty are the most rigorous methods of cost risk analysis. What is needed is a systematic set of empirical case studies of elicitation in cost risk analysis to allow retrospective studies of the effectiveness and accuracy of different techniques. These case studies will provide cost risk analysts with a set of credible tools to do elicitation that can be compared and refined with further experience. Cost risk analysis is in a unique position to contribute to the development of elicitation procedures: It has a need for elicitation to quantify significant uncertainties, it has many different opportunities in government and industry to apply these techniques and test them, and it has quantitatively sophisticated practitioners who can help advance the field of elicitation.

---

upper limit. The value of the density at the most-likely value is chosen so that the density integrates to one, as required for a probability density.

Acknowledgments

The author would like to thank the following colleagues for thoughtful comments on earlier drafts of this report: Mark V. Arena, Susan M. Paddock, Giles K. Smith, Don Snyder, Fred Timson, and Obaid Younossi at RAND; Joe Hamaker, Director, NASA Headquarters Cost Analysis Division; and James S. Hodges at the University of Minnesota. RAND colleagues Robert Leonard and J. R. Lockwood provided thoughtful and constructive reviews. Any remaining errors are the responsibility of the author.

Abbreviations

CER	cost-estimating relationship
NASA	National Aeronautics and Space Administration
PAF	Project AIR FORCE
SSCAG	Space Systems Cost Analysis Group
WBS	Work Breakdown Structure

CHAPTER ONE

Introduction

It has become a truism of cost estimation that there is no "right" cost estimate: For any particular estimate, there is some probability that the actual cost will be higher or lower than that estimate. A cost estimate is only a forecast. It has inherent uncertainties because requirements, technology, the economic environment, political considerations, and a multitude of other factors are likely to change before the system whose cost is being estimated has been assembled.

Recognition of these uncertainties has led cost-estimation practitioners to advocate that cost estimates be expressed not as a single point estimate but as a range of "likely" costs. Starting with Sobel (1965) and Dienemann (1966) through such researchers as Book (2001) and Garvey (2000) today, one major approach to quantifying cost uncertainty has been probabilistic: expressing the uncertainty in a cost estimate as a probability distribution and using that distribution to compute such quantities as expected cost and most-likely cost and to help set budget levels with various probabilities of overrun. This approach is advocated in a variety of publications, including the National Aeronautics and Space Administration's (NASA's) *NASA Cost Estimating Handbook* (2004), the *Risk Management Guide for DoD Acquisition* (Department of Defense, 2003), a book-length treatment by Paul Garvey (2000), and numerous articles and tutorials by other prominent workers in the field.[1]

Early papers on this approach simply assumed that such a probability distribution for cost had already been constructed, and they proceeded to explicate what could be done with the distribution to inform decisionmakers. But it soon became evident that the actual construction of such a probability distribution for cost was not a trivial exercise. In a review of cost uncertainty methodology, Fisher (1962) praised the analytic rigor of the probabilistic approach, but he noted that it was difficult to construct the distributions while expressing optimism that, eventually, techniques could be developed to get the underlying distributions.

The use of *cost-estimating relationships* (CERs), which are usually regression equations based on historical data, provides a measure of uncertainty of a cost estimate as a confidence interval for prediction. However, that estimate assumes that the values of the independent variables of the CER are known exactly, which is not true for a project in the early stages of planning. Further, historical data from which to construct credible CERs are sometimes not available—as when a technology or manufacturing process is radically new (see, e.g., Kitchenham et al., 2002). To get probability distributions for values of independent variables in CERs or for the characteristics of new and untried technologies, cost analysts have often proposed tapping

---

1 Other approaches to describing and reasoning with uncertainty take advantage of expert opinion, but they do not require formal probability distributions, although such distributions can be used. Two notable examples are the Delphi method, developed at RAND in the 1950s and extended and widely applied since then to summarize views of expert panels (see, e.g., Adler and Ziglio, 1995), and assumption-based planning, which aims to relate potential paths of a project to specific sets of alternative assumptions (Dewar, 2002).

the resources of expert judgment, such as that possessed by engineers, managers, and other knowledgeable people, and constructing, or eliciting, subjective[2] probability distributions to represent uncertainty. This process is called *elicitation* in the wider literature of decision analysis, and it is known to be difficult to execute and can be subject to numerous biases. However, its potential utility and the fact that it is widely advocated in cost risk analysis require that we understand how to do it well in order to have credible risk analyses.[3]

This report is based on a selective review of the literature of elicitation, both in the cost risk field and in other areas in which elicitation has been a topic of research, primarily statistics and psychology. Because of a lack of empirical work on elicitation, especially cost risk, the author also interviewed a number of senior people in the cost risk community who gave insight into the practices of the field.

---

2 The probability distributions so determined are termed *subjective* because they represent the personal opinion of an expert. *Objective* probability distributions would be those determined from measurements on a physical system.

3 Some authors argue that expert opinion should be used very sparingly, if at all (see, e.g., Conrow, 2003).

CHAPTER TWO

Elicitation in Decision Analysis

Elicitation of subjective probability distributions as an area of research in its own right arose out of several developments in the 1950s and 1960s. The first was a renewed interest in Bayesian statistics. Practitioners of Bayesian statistics argued that probability should reflect a subjective state of knowledge in which a rational person would use Bayes' theorem to modify an initial state of knowledge about some item of interest, such as a final cost (the prior probability distribution), with acquired data to form an updated probability distribution (a posterior probability distribution). This argument contrasted with the prevailing school of frequentist statistics, which held that the only basis for assigning probability to an event was its chance of occurring in a large run of experiments.[1] Bayesian statistics, therefore, requires two steps: eliciting the prior distribution and then performing the mathematical calculations required to apply Bayes' theorem.

The parallel development of general risk analysis, with its need to quantify probabilities of hazardous events that might be either extremely rare or stemming from some new and untried technology, was also an impetus for developing many of the more sophisticated elicitation methods. For example, a major study on nuclear-reactor safety used extensive elicitation and culminated in a guide to elicitation methods (Wheeler et al., 1989). These methods were later used and extended in the area of environmental risk, which required quantifying the prevalence of environmental hazards, the exposures of various populations, and the effects of those exposures on individuals' health, all questions that often had scant empirical data. For a detailed review and examples, see Morgan and Henrion (1990).

As a result of this interest, a number of psychologists began to study the general topic of human decisionmaking under uncertainty. In a typical experiment, subjects would be asked questions whose answers were unknown to them, and then would be asked to quantify their uncertainty about their answers. Usually, this uncertainty was expressed in terms of a probability distribution by the researchers, although the questions asked of the subject may or may not have used probabilistic terminology. These experiments indicated that human beings were

---

1 The differences between the Bayesian and frequentist approaches to statistics are intensely debated in the field of statistics. The debate stems from some significant differences between the two approaches in the methods of statistical inference that are held to be valid. Advocates of the Bayesian approach, such as Savage (1954) and Lindley (1983) have argued that it is the only rational way to perform statistical inference and have arrayed evidence to show that the frequentist approach could lead to absurd decisions, even in relatively simple cases. A more detailed description of these differences is tangential to the discussion here.

subject to a number of biases that distorted their judgment about the uncertainty of their knowledge. The most commonly listed biases encountered were[2]

- availability: the tendency to overestimate the probability of events that are easy to recall
- representativeness: judging the probability of events by focusing on characteristics (possibly irrelevant) in which they resemble other events
- anchoring and adjustment: the bias of the final assessment of a value toward an initial assessment of the value by constraining subsequent adjustment of probabilities in the light of new evidence
- overconfidence: underestimation of uncertainty about a quantity.

However, there were a number of criticisms of this research that are relevant for elicitation practice. First, the vast majority of the experiments were done with subjects (typically university students) who were not experts in the areas in which they were being questioned[3] and who were generally not familiar with probability concepts. In several attempts to carefully study the elicitation of truly expert opinion, the results were mixed. Some researchers found that experts' judgments in their own fields were not subject to one or more of the common biases.[4] Other researchers found that experts' performance worsened when they were asked questions in areas in which they were not expert (Mullin, 1986), but there were still other researchers who found that expert judgment in qualitative fields was suspect (Tetlock, 2005).

A further criticism by Edwards (1975) noted that the testing situation itself was artificial, since it typically denied the experimental subjects, whether novice or expert, the use of intellectual tools, such as reference materials or computational devices, that would normally be available to them.[5] Other studies looked at the effect of decomposing elicitation tasks into judgments of probabilities of simpler events. Some authors' results showed that decomposition improved performance; those of others indicated problems with decomposition and uncertainty judgment. Later reviewers of this work have equivocated (Morgan and Henrion, 1990).

The relevance of the psychological research for expert elicitation is in some doubt, because little empirical research attempted to tie the research to elicitation practice, a fact noted by several researchers who attempted to synthesize the literature (Morgan and Henrion, 1990; Meyer and Booker, 2001; and Garthwaite, Kadane, and O'Hagan, 2004). In the case of biases, for example, only two suggestions were made: (1) Have an elicitor ask the subject to explicitly think of reasons why his or her initial estimates might not be correct (to counter overoptimism and anchoring) and (2) ask for range extremes first, instead of the most likely value. How-

---

[2] The canonical list is in Kahneman, Slovic, and Tversky (1982), but variants are given in Mullin (1986); Hogarth (1989); Morgan and Henrion (1990); Wolfson (1995); and Garthwaite, Kadane, and O'Hagan (2004).

[3] The questions used in these studies were often simple, factual questions, such as the distance between two cities, and were therefore often termed "almanac questions."

[4] Weather forecasters are particularly good at probability judgments when forecasting weather conditions (see Morgan and Henrion, 1990, p. 130).

[5] Note that the elicitations described in Mullin and in Morgan and Henrion did, in fact, allow the subjects complete access to these materials.

ever, little empirical evidence documented the effect of these suggestions, as plausible as they seemed.

After the initial interest in the 1960s and 1970s, development and interest in these issues lagged in the statistical field until the mid-1990s, because the mathematical computations required to compute posterior distributions and implement Bayesian inference were intractable, requiring numerical integration of complex multidimensional integrals, except in certain special cases. As a result, Bayesian statisticians tended to confine their attention to simple standard prior-probability distributions, not dealing with elicitation.[6] However, in general risk analysis, interest in and application of elicitation methods persisted.

Then, in the late 1980s, a new computational tool, Markov Chain Monte Carlo simulation, allowed Bayesian statisticians to use the rapid increase of computational power to accurately estimate posterior distributions for essentially arbitrary prior distributions, including those developed using elicitation. Interest has continued to the present time, as indicated by a number of recent publications addressing issues of practical elicitation (Chaloner, 1996; Wolfson, 1995; Kadane and Wolfson, 1998; O'Hagan, 1998; and Meyer and Booker, 2001). The result is a substantial body of work on elicitation on which the cost risk analysis field can draw.

---

[6] There were exceptions (see, e.g., Kadane, 1996).

CHAPTER THREE

Elicitation in Cost Risk Analysis

Elicitation in cost risk analysis[1] focuses on obtaining a subjective cost probability distribution directly or (more commonly) eliciting a subjective probability distribution for some project characteristic that is a cost driver, such as weight, power usage, or development schedule. Since these variables are used as independent cost drivers in CERs, their subjective distributions can be used to compute a predictive distribution for cost that includes uncertainties in the inputs as well as in the estimating relationship. The resulting distributions for subsystems can be added to other cost distributions via Monte Carlo simulation or analytic methods to get an overall cost probability distribution for the entire project (Garvey, 2000; Arena et al., 2006).

The actual methods of elicitation for cost risk purposes are somewhat difficult to determine, because there is little information in the professional literature, other than tutorials, that actually documents how elicitation is done. The tutorials generally recommend asking an expert for the maximum, minimum, and most-likely values of the quantity whose distribution is being elicited, and then fitting a triangle distribution[2] to the three numbers (Morgan and Henrion, 1990, or Garvey, 2000). Some tutorials may recommend asking an expert for percentiles[3] of the distribution and then fitting a normal, a log normal, or a beta distribution to these quantities, but specific information on how this fitting should be done, how the percentiles

---

[1] In formal decision analysis, a distinction is made between uncertainty and utility. *Uncertainty* refers to the probability of an event (usually something untoward) occurring, whereas *utility* measures the consequences of the event to a decisionmaker. A formal risk analysis, therefore, combines both probability and utility. The distinction is made because consequences are rarely linear (a $5-million cost overrun may be much more than five times as problematic to a decisionmaker as a $1-million dollar overrun, because of reporting requirements, oversight, etc. However, in discussing cost risk, analysts rarely take the added step of determining utility. Therefore, the term *cost risk* is used where *cost uncertainty* would be a more precise description from a formal decision-analysis point of view. See DeGroot (1970) for a clear exposition of utility in decision analysis.

[2] For the triangle distribution, the probability is set to zero outside the endpoints, while between the endpoints the density rises linearly from the lower value to the most-likely values and then decreases linearly from the most-likely value to the upper limit. The value of the density at the most likely value is chosen so that the density integrates to one, as required for a probability density.

[3] The *X*th *percentile* is defined as the point in a probability distribution where *X* percent of the probability lies below that point. For example, when an expert specifies the 90th percentile of a cost probability distribution, this would indicate a belief that 90 percent of the time, the cost would be less than that number.

should be checked, and how the methods perform with respect to known biases in elicitation is largely absent in the cost risk literature.[4]

However, elicitation has been treated in several other areas of decision analysis, so this literature can be surveyed to evaluate elicitation practices in cost risk analysis.

---

[4] For example, Lurie, Goldberg, and Robinson (1993) emphasize the mathematical and probability aspects of cost risk analysis and assume that the distributions have already been determined.

CHAPTER FOUR

Current Best Practices

What practical advice can be given? While there seems to be much to consider and change in elicitation for cost risk analysis, a start can be made with the following procedure, which is synthesized from a number of sources, including current practice in cost risk (Morgan and Henrion, 1990; Chaloner, 1996; and Meyer and Booker, 2001):

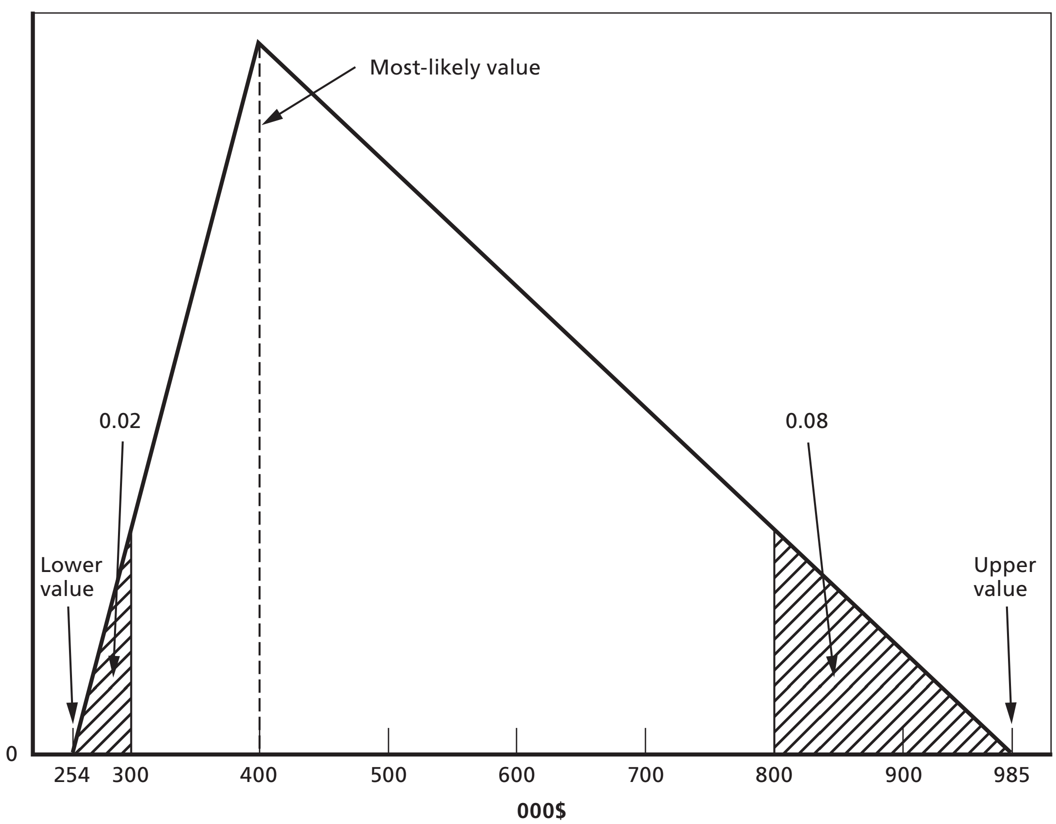
- Use multiple experts, if possible. If program engineers from the project whose cost is to be estimated are used for some of the elicitations, independent engineers should also be included if at all feasible.
- Ask the expert to provide, at a minimum, upper, lower, and most-likely values for cost of the Work Breakdown Structure (WBS) element under consideration (or for the technical characteristic that drives cost). During the elicitation, the expert should be pushed to think of reasons that the range could be larger (especially in the upper direction), and to explain the reasoning behind the answers. This extension will counteract tendencies to overoptimistically narrow distributions and will give the expert and the person conducting the elicitation insight into issues that might be useful in further elicitation or analysis. And, although many cost risk analysts report that they ask for the most-likely value first, the literature suggests that the central value should be elicited near the end of the elicitation to help counteract any effect of anchoring (Morgan and Henrion, 1990, p. 149; Spetzler and Von Holstein, 1975).
- Fit a triangle distribution to the three numbers elicited, but use the upper and lower values to bound 90 percent of the probability (where reasonable).[1] Some authors refer to this bounded version as the "expanded" triangle, which adds more spread to the distribution and helps to counteract overoptimism. See Figure 4.1 for a notional example in which elicitation gave 300, 400, and 800 for the lower, most-likely, and upper values. Using Garvey's procedure for distributing the remaining 10 percent, we get a triangle distribution with 254 for the lower value and 985 for the upper value. Note that 10 percent of the probability is not distributed equally in the tails of the distribution: Two percent is

[1] The triangle distribution is often used because of its simplicity—e.g., Morgan and Henrion (1990) and Book (2001). The extension of the endpoints seems to be part of the folklore of practice: Garvey (2000) gives some convenient formulas (used here) that distribute the remaining probability between the two tails in proportion to the skewness of the elicited upper and lower values. Biery, Hudak, and Gupta (1994) recommend a variant that divides the remaining probability equally between the two tails.

in the lower tail and 8 percent is in the upper one. This skewness is driven by the location of the most-likely value with respect to the original upper and lower values.

- Most current authors recommend eliciting at least two more percentiles of the risk distribution and see if they are consistent with the fitted triangle or expanded triangle distribution. The new percentiles will therefore provide a valuable check on the elicitation. Also, it is recommended that percentiles be elicited in multiple ways to help check and diagnose bias.
- Provide feedback to the expert about the results of the elicitation, preferably in the same elicitation session. Such feedback would include the display of newly elicited percentiles described above, but might also include the final range of nonzero probabilities, the median estimated cost, the probability that the final cost will exceed the most-likely cost, and so on. It would be helpful to also be able to display historical data, if available.
- Carefully document the elicitation process, describe the results obtained, and archive the data for future retrospective studies.

Figure 4.1
Fitting an Expanded Triangle Distribution to Upper, Lower, and Most-Likely Values



RAND *TR410-4.1*

CHAPTER FIVE

Elicitation in Cost Analysis

How do the elicitation methods used in the cost risk community compare with the suggested methodologies and limited empirical studies in the psychological and statistical literature? As noted previously, the initial literature of cost risk analysis displayed little interest in the practicalities of elicitation, even while routinely recommending elicitation of expert judgment when data were scarce or when historical data might be irrelevant. In addition, the open cost analysis literature has had few articles on techniques. Perhaps most surprising, a review shows little overlap in the literatures of elicitation in cost risk analysis with that of elicitation in other fields, such as general risk analysis, statistics, and psychology.

In the mid-1980s, Wallenius pointed out that a key review paper on the current state of cost estimation had no overlap in citations with the book on judgment under uncertainty by Kahneman, Slovic, and Tversky (Wallenius, 1985, referring to Kahneman, Slovic, and Tversky, 1982; and McNichols, 1984). This lack of overlap has largely continued until today: In general, when cost analysis authors do touch on elicitation and reference any sources outside the cost analysis field, they point to the most recent major review of uncertainty and refer readers to it for more detail on how to do an elicitation.

Until the early 1990s, the preferred reference was Kahneman, Slovic, and Tversky, which actually did not provide a good set of practical guidelines for elicitation, although it documented the biases to which elicitors (particularly naive ones) were subject. Since then, the preferred reference has been the book by Morgan and Henrion (1990), which does, in fact, provide considerable guidance on procedure.

However, as with the general elicitation literature, there is little discussion in the open cost risk literature of the elicitation processes actually used in actual projects. In most publications, elicitation of probability is given shorter shrift than are calculations and final results.

But even the sketchy details that are given indicate that there are important gaps between practice in the cost analysis field and the practice recommended by Morgan and Henrion and Meyer and Booker. Meyer and Booker devote much time and energy to selecting experts, preparing initial written materials for the experts on the problem and its context, and doing the elicitation. At least some of the procedures are designed to counteract the classic elicitation biases enumerated above, and, in all cases, care is given to feeding back the results of the elicitation to the experts in a form that allows the experts to see the implications of their judgments and perhaps revise the quantification of their beliefs. Perhaps most important, these authors recommend carefully documenting the elicitation procedures and results.

In addition, elicitation practices in cost risk analysis are very diverse, with little standardization, even in areas such as cost estimation for acquisition of space systems, in which one might expect a convergence of practice, given that such acquisition decisions are concentrated in a few government agencies. Individual organizations also vary greatly. However, based on interviews conducted by the author,[1] a number of worrisome issues are common in current elicitation practice in the cost risk field:

- Experts selected for elicitation should have technical expertise to understand technological issues, management experience to appreciate the organizational challenges that can arise, and, above all, be independent of the project under review. However, in practice, selection of experts is often a matter of convenience of access. In many cases, the only readily available experts are engineers from the program office of the project whose cost is being estimated, and these engineers provided the initial technical and/or cost estimates; they tend to produce a very narrow distribution around their initial estimate.
- Elicitation is often rushed due to time constraints of the experts and time and financial constraints on the team doing the elicitation—particularly in some of the cost analysis organizations in the Department of Defense, where there have been significant staff cuts over the past decade. In many cases, the elicitation is done by mail or Web form, with little interaction with the subject.
- Feedback is rarely given to the expert about the implications of the elicitation, even in terms of historical data.
- The elicitation methodologies are largely ad hoc, in that they are seldom based on or derived from references to the elicitation literature or verified by internal historical assessments of effectiveness.
- Little or no documentation is prepared for or retained about the process, forms used, and so on. It is especially hard to go back to finished projects and get historical information about the elicitations that were done.
- As a consequence of the last point, it is almost impossible to go back over elicitations and do an analysis of how accurate they were in capturing the final costs.
- As a consequence, it is difficult to give advice to cost risk analysts contemplating elicitation on how to allocate scarce resources.

There are some special characteristics of cost risk analysis that might justify modifications to elicitation practices in other fields. For example, cost risk analysts typically have to elicit many distributions in the course of a risk analysis for a complex project. The cost risk literature recommends doing a cost risk simulation using numbers of project elements (typically enu-

---

[1] The interviewees included cost-estimation professionals at the U.S. Air Force's Cost Analysis Agency, the Air Force's Space and Missile Center, NASA Headquarters, the Jet Propulsion Laboratory, and several private aerospace and consulting firms. The points that follow are the author's synthesis of these discussions; as such, they are limited by the range of people interviewed. However, the sources included several senior figures who have written and consulted widely in cost risk analysis and who are in a position to comment on the practices in the field. Their comments are also consistent with the literature reviewed.

merated in the WBS, which are in the high tens to low hundreds.[2] In comparison, the outside elicitation literature typically works with many fewer elicitations. Documenting and archiving elicitation materials costs money, and there are currently no sources for such funds, at least in government organizations, absent the interest and direction of senior leadership.

Interviews with cost risk analysts in private aerospace and consulting firms suggest that these organizations do a somewhat more careful job of elicitation. However, their elicitation practices are considered proprietary and the organizations are reluctant to describe what they do in detail for public disclosure.

---

[2] Cost risk tools, such as ACEIT from Tecolote Research, can handle thousands of individual cost elements, probably too many to elicit individually.

CHAPTER SIX

Conclusions

The cost-estimation community is in general agreement that probabilistic methods of quantifying and reasoning with uncertainty are the most rigorous methods of cost risk analysis. These methods may not always be used, either because not enough time and resources are available or because of the detail required for a particular purpose (Arena et al., 2006). When relevant historical data are not available, however, elicitation of expert opinion is acknowledged to be a reasonable alternative. But, although a set of procedures for careful, documented, controlled elicitation is emerging that attempts to deal with known biases (Morgan and Henrion, 1990; Meyer and Booker, 2001; and Garthwaite, Kadane, and O'Hagan, 2004), it is fair to say that these procedures are not followed generally in the cost risk community, based on the interviews and the available public literature. Further, to date there has been little comment on or explanation of this gap in the community.

This is not to say that elicitation research outside of the cost risk community has a definitive set of answers. The actual performance of elicitation procedures designed to minimize the classic biases of anchoring, optimism, and so on, has not been studied extensively (see Mullin, 1986, for a partial example), and there may well be enhancements that are necessary to achieve the de-biasing required for more-accurate assessments of uncertainty. There is also some evidence of substantial differences in the uncertainty judgments of expert as opposed to naive subjects, which means that some of the biases that have concerned researchers in the past may not be relevant to elicitation in cost risk.

However, elicitation practice in cost risk analysis needs to be improved significantly before it should begin to be concerned about issues such as these. There are a number of steps the cost risk community could take to improve its use of elicitation:

- Assemble a reasonably complete list of current elicitation practices in cost risk analysis. This list would include, for example, the "expanded triangle," but other methods may be in use that have not been described in the literature.
- The methods should be critically examined for their theoretical and empirical underpinnings, using the wider elicitation literature, or they may be used by other disciplines.
- The performance of these methods in eliciting cost and the other uncertainties used in cost risk analysis should be tracked with empirical case studies and a database of elicited distributions and actual costs that occurred. And there should be enough documentation to allow retrospective studies. Standards should be set for documenting the application of

each of the methods to make it easier to assemble evidence for a method's strengths and drawbacks in particular situations.

These steps should provide cost risk analysts with a set of credible tools to do elicitation that can be compared and refined with further experience. The professional groups and major meetings in cost estimation (e.g., the Space Systems Cost Analysis Group, International Society for Parametric Analysis, the annual Department of Defense Cost Analysis Symposium) and the cost-estimation journals should encourage the publication of such research, both theoretical and actual case studies, and should insist that the reporting of the use of elicitation be accompanied by information about the process used.

Finally, long-term studies of the performance of different methods in capturing uncertainties should be carried out by comparing elicited distributions with later actual costs. A long line of articles in the literature has consistently noted this key lack,[1] and virtually all cost analysts interviewed by the author agreed. Arguments against this endeavor include expense, lack of time in understaffed organizations, the long time frames involved, and the unavoidable changes in projects. All of these factors make comparisons difficult, but without such comparisons, how can the value of elicitation be judged? The field is left with anecdotes or, worse, the suspicion that elicitation is only a crutch to get a set of required numbers at the end of the process that have no real value in helping to make decisions. Hilson (1998), commenting on project risk management, made this point explicitly:

> In the absence of a coherent body of irrefutable evidence, the undoubted benefits that can accrue from effective management of risk must currently be taken on trust. Overcoming this will require generation of a body of evidence to support the use of formal project risk management, providing evidence that benefits can be expected and achieved, and convincing the skeptical or inexperienced that they should use project risk management.

Some of this information may be considered to be proprietary by commercial firms, notwithstanding their participation in and support of professional societies. However, government agencies, such as the Department of Defense and NASA, have no such constraints and do have an interest in ensuring that the best procedures are available for all to use.

Finally, the cost risk field (and cost estimation in general) would be well served by using and citing relevant literature in other fields, such as statistics and psychology. In addition, the cost risk literature could be made more accessible outside the field: The literature largely appears in specialized and sometimes short-lived journals, or in conference proceedings that are difficult to access just a few years after publication. Collecting the literature and making it more easily available might be a worthwhile project for the professional societies supported by government and industry.

Cost risk analysis is in a unique position to contribute to the development of elicitation procedures: It has a need for elicitation to quantify significant uncertainties that affect impor-

---

[1] Beach (1975); discussions of papers on "Elicitation" (1998); Meyer and Booker (1991); Morgan and Henrion (1990); O'Hagan (1998); Hilson (1998); and Kitchenham et al. (2002).

tant decisions, it has many different opportunities in government and industry to apply these techniques and test them, and it has quantitatively sophisticated practitioners who can help advance the field of elicitation. But to do so, it has to take elicitation seriously and upgrade the techniques used across the profession.

Bibliography

Adler, Michael, and Erio Ziglio, *Gazing into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health*, London, UK: Jessica Kingsley Publishers, 1995.

Arena, Mark V., Obaid Younossi, Lionel A. Galway, Bernard Fox, John C. Graser, Jerry M. Sollinger, Felicia Wu, and Carolyn Wong, *Impossible Certainty: Cost Risk Analysis for Air Force Systems*, Santa Monica, Calif.: RAND Corporation, MG-415-AF, 2006. As of September 26, 2006:
http://www.rand.org/pubs/monographs/MG415/

Beach, Barbara H., "Expert Judgment About Uncertainty: Bayesian Decision Making in Realistic Settings," *Organizational Behavior and Human Performance*, Vol. 14, 1975, pp. 10–59.

Book, Stephen A., "Estimating Probable System Cost," *Crosslink*, Winter 2001, pp. 12–21. As of September 26, 2006:
http://www.aero.org/publications/crosslink/winter2001/

Biery, Fred, David Hudak, and Shishu Gupta, "Improving Cost Risk Analyses," *Journal of Cost Analysis*, Spring 1994, pp. 57–85.

Chaloner, Kathryn, "Elicitation of Prior Distributions," in Donald A. Berry and Dalene K. Stangl, eds., *Bayesian Biostatistics*, New York: Marcel Dekker, 1996.

Conrow, Edmund H., *Effective Risk Management: Some Keys to Success*, Reston, Va.: American Institute of Aeronautics and Astronautics, 2003.

DeGroot, Morris H., *Optimal Statistical Decisions*, New York: McGraw-Hill, 1970.

Dewar, James A., *Assumption-Based Planning: A Tool for Reducing Avoidable Surprises*, New York: Cambridge University Press, 2002.

Diekemann, James E., and W. David Featherman, "Assessing Cost Uncertainty: Lessons from Environmental Restoration Projects," *Journal of Construction Engineering and Management*, November/December 1998, pp. 445–451.

Dienemann, Paul F., *Estimating Cost Uncertainty Using Monte Carlo Techniques*, Santa Monica, Calif.: RAND Corporation, RM-4854-PR, 1966.

Discussions of papers on "Elicitation," in *The Statistician*, Vol. 47 (Part 1), 1998, pp. 55–68.

Edwards, Ward, "Comment" [on Hogarth (1975)], *Journal of the American Statistical Association*, Vol. 70, No. 350, 1975, pp. 291–293.

Fisher, G. H., *A Discussion of Uncertainty in Cost Analysis*, Santa Monica, Calif.: RAND Corporation, RM-3071-PR, 1962.

Galway, Lionel A., *Quantitative Risk Analysis for Project Management: A Critical Review*, Santa Monica, Calif.: RAND Corporation, WR-112-RC, 2004. As of September 26, 2006:
http://www.rand.org/pubs/working_papers/WR112/

Garthwaite, Paul H., Joseph B. Kadane, and Anthony O'Hagan, *Elicitation*, Pittsburgh, Pa.: Carnegie Mellon University, Department of Statistics, Technical Report 808, 2004.

Garvey, Paul R., *Probability Methods for Cost Uncertainty Analysis*, New York: Marcel Dekker, 2000.

Hilson, David, "Project Risk Management: Future Developments," *International Journal of Project and Business Risk Management*, 1998.

Hogarth, Robin M., "Cognitive Processes and the Assessment of Subjective Probability Distributions," *Journal of the American Statistical Association*, Vol. 70, No. 350, 1975, pp. 271–294.

Hogarth, Robin M., *Judgment and Choice*, New York: John Wiley, 1989.

Kadane, Joseph B., ed., *Bayesian Methods and Ethics in a Clinical Trial Design*, New York: Wiley & Sons, 1996.

Kadane, Joseph B., and Lara J. Wolfson, "Experiences in Elicitation," *The Statistician*, Vol. 47, No. 1, 1998, pp. 3–19.

Kahneman, Daniel, Paul Slovic, and Amos Tversky, *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge, UK: Cambridge University Press, 1982.

Kitchenham, Barbara, Shari Lawrence Pfleeger, Beth McColl, and Suzanne Eagan, "An Empirical Study of Maintenance and Development Estimation Accuracy," *Journal of Systems and Software*, Vol. 64, No. 1, October 15, 2002, pp. 57–77.

Lindley, Dennis V., "Theory and Practice of Bayesian Statistics," *The Statistician*, Vol. 32, 1983, pp. 1–11.

Lurie, Philip M., Matthew S. Goldberg, and Mitchell S. Robinson, *A Handbook of Cost Risk Analysis Methods*, Alexandria, Va.: Institute for Defense Analyses, P-2734, 1993.

McNichols, Gerald R., "The State-of-the-Art of Cost Uncertainty Analysis," *Journal of Cost Analysis*, Vol. 1, 1984, pp. 149–174.

Meyer, Mary A., and Jane M. Booker, *Eliciting and Analyzing Expert Judgment: A Practical Guide*, Philadelphia, Pa.: Society for Industrial and Applied Mathematics and the American Statistical Association, 2001.

Morgan, Millett Granger, and Max Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, New York: Cambridge University Press, 1990.

Mullin, Theresa M., *Understanding and Supporting the Process of Probabilistic Estimation*, Ph.D. thesis, Carnegie Mellon University, School of Urban and Public Affairs, 1986.

National Aeronautics and Space Administration, *NASA Cost Estimating Handbook 2004*, Washington, D.C., 2004. As of September 26, 2006:
http://www.nasa.gov/offices/pae/references/index.html

O'Hagan, Anthony, "Eliciting Expert Beliefs in Substantial Practical Applications," *The Statistician*, Vol. 47, No. 1, 1998, pp. 21–35.

O'Hagan, Anthony, and Jeremy E. Oakley, "Probability Is Perfect, but We Can't Elicit It Perfectly," *Reliability Engineering and System Safety*, Vol. 85, 2004, pp. 239–248.

Raymond, Fred, "Quantify Risk to Manage Cost and Schedule," *Acquisition Review Quarterly*, Spring 1999, pp. 147–155.

Savage, Leonard J., *The Foundations of Statistics*, New York: Dover Publications, 1972 (reprint of 1954 edition).

Sobel, S., *A Computerized Technique to Express Uncertainty in Advanced System Cost Estimates*, Bedford, Mass: MITRE, ESD-TR-65-79, 1965.

Spetzler, Carl S., and Carl-Axel S. Von Holstein, "Probability Encoding in Decision Analysis," *Management Science*, Vol. 22, No. 3, 1975, pp. 340–358.

Tetlock, Philip E., *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, N.J.: Princeton University Press, 2005.

U.S. Department of Defense, *Risk Management Guide for DoD Acquisition*, 5th ed., Ft. Belvoir, Va.: Defense Acquisition University Press, 2003.

Wallenius, K. T., "Cost Uncertainty Assessment Methodology: A Critical Overview," *Department of Defense Cost Analysis Symposium Proceedings*, Washington, D.C.: U.S. Department of Defense, 1985.

Wheeler, T. A., S. C. Hora, W. R. Cramond, and S. D. Unwin, *Analysis of Core Damage Frequency from Internal Events: Expert Judgment Elicitation*, Vol. 2, Washington, D.C.: Nuclear Regulatory Commission, NUREG/CR-4550, 1989.

Wolfson, Lara J., *Elicitation of Priors and Utilities for Bayesian Analysis*, Ph.D. thesis, Carnegie Mellon University, Department of Statistics, Pittsburgh, Pa., 1995.